National Library of Canada
Bibliothèque nationale du Canada

Canadian Theses Service
Services des thèses canadiennes

Ottawa, Canada
K1A 0N4

CANADIAN THESES
THÈSES CANADIENNES

NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30.

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30.

THIS DISSERTATION
HAS BEEN MICROFILMED
EXACTLY AS RECEIVED

LA THÈSE A ÉTÉ
MICROFILMÉE TELLE QUE
NOUS L'AVONS REÇUE

Canada

NL-339 (r.86/06)

THE UNIVERSITY OF ALBERTA

Memory for typical and atypical elements in an aggressive event

BY

© Sara Elizabeth Comish

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF ARTS

DEPARTMENT OF PSYCHOLOGY

EDMONTON, ALBERTA

FALL, 1986

Permission has been granted to the National Library of Canada to microfilm this thesis and to lend or sell copies of the film.

The author (copyright owner) has reserved other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without his/her written permission.

L'autorisation a été accordée à la Bibliothèque nationale du Canada de microfilmer cette thèse et de prêter ou de vendre des exemplaires du film.

L'auteur (titulaire du droit d'auteur) se réserve les autres droits de publication; ni la thèse ni de longs extraits de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation écrite.

ISBN 0-315-32447-3

THE UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR: Sara Elizabeth Comish

TITLE OF THESIS: Memory for typical and atypical script elements in an aggressive event

DEGREE: Master of Arts

YEAR THIS DEGREE GRANTED: 1986

Permission is hereby granted to THE UNIVERSITY OF ALBERTA LIBRARY to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

S. E. Comish

(Student's signature)

#309 715 VANCOUVER ST

(Student's permanent address)

VICTORIA B.C. V8D 3V2

Date: AUG 26/1986

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled MEMORY FOR TYPICAL AND ATYPICAL SCRIPT ELEMENTS IN AN AGGRESSIVE EVENT submitted by SARA ELIZABETH COMISH in partial fulfilment of the requirements for the degree of MASTER OF ARTS

Brendan Gail Rule

(Supervisor)

Allen R Dobbs

J Paul Johnston

Date: AUG 8/1986

FOR JOHN

(maybe now you'll know what i do)

Abstract

Scripts are knowledge structures about ordered events. Script theorists have proposed that script structure enables typical, script elements to be processed differently from atypical, non-script elements and further, that the processing differences result in differential retrieval of typical and atypical elements. The first goal of this thesis was to examine whether the script was used in the same way in recall as it was used in recognition. This was accomplished by comparing the pattern of hits and false alarms for typical and atypical items in recall to the pattern of hits and false alarms in recognition. One hundred and twenty introductory psychology students heard a story about an evening at a restaurant and heard a story about a marital dispute. Later they were asked to do either a recall or a recognition task for each story. The pattern of hits and false alarms was the same for recall as the pattern for recognition. This finding was consistent both for the restaurant story and for the affect-laden marital dispute story suggesting that script research can be extended to affect-laden material. The second goal of this thesis was to examine script content. It was determined, by examining the intrusions in recognition of 60 introductory psychology students, that emotional states, emotional expressions, and physical aggression are part of people's knowledge structures concerning marital disputes.

ACKNOWLEDGEMENT

I wish to thank my supervisor, Brendan Rule, for all her kind help with the thesis. In addition, I am grateful to the other members of my committee, Allen Dobbs, and Paul Johnston, for their helpful comments.

Plus, I'm very grateful to Allison Matichuk and Susan Boon for their generous help with the research and to John Turtle, a well-known graphic artist.

(And of course to my parents, Keith and Betty Comish)

TABLE OF CONTENTS

CHAPTER		PAGE
I.	INTRODUCTION	1
	Background to the issue	1
II.	STUDY ONE	10
	Pretest	10
	Rationale	11
	Method	11
	Results	16
III.	STUDY TWO	29
	Rationale	29
	Method	30
	Results	31
IV.	GENERAL DISCUSSION	38

REFERENCES....44

Appendix A. Pilot study ratings....47

Appendix B. Stimulus materials....50

Appendix C. Phony questionnaire....55

Appendix D. A sample recall....56

Appendix E. Sample recalls....57

LIST OF TABLES

Table	Description	Page
1	The percentage of hits and false alarms for the restaurant story with contrast weights	18
2	The percentage of hits and false alarms for the marital dispute story with contrast weights	20
3	The percentage of hits and false alarms by men and women	22
4	The percentage of hits and false alarms by retrieval	24
5	Memory scores for typical and atypical items in recall and recognition over time	27
6	Frequencies of "yes" responses to one or more emotional states, emotional expression, physical aggression	33
7	Frequencies of "yes" responses for typical and atypical emotional states, emotional expression, physical aggressions	35
8	Frequencies of false alarm responses for typical emotional states, emotional expression, physical aggressions	37

LIST OF FIGURES

Figure		Page
1a.	Recognition hits and false alarms data from Smith and Graesser (1981)	58
1b.	Recall hits and false alarms data from Smith and Graesser (1981)	60
2.	Predicted pattern of hits and false alarms for recall and recognition	62
3.	The percentage of hits and false alarms for the restaurant/recall conditions	64
4.	The percentage of hits and false alarms for the restaurant/recognition conditions	66
5.	The percentage of hits and false alarms for the marital dispute/recall conditions	68
6.	The percentage of hits and false alarms for the marital dispute/recognition conditions	70

Because we tend to make order out of the "blooming, buzzing confusion" of our world, events in our world appear organized and consistent. One of the contemporary theoretical approaches in social cognition that addresses the organization of events is schema theory. Schema theory suggests that memory is guided by a knowledge structure that guides the processing and modification of incoming information. This structure is known as a schema. The schema concept was originally applied to the phenomenon of story memory by Bartlett (1932) in order to account for distortions that he had observed in story recall. He believed that schemata were involved in the active organization of past experiences and operated at retrieval (Bartlett, 1932, pp.206).

Scripts are related to schemata. Scripts are knowledge structures about events in which actions occur in a certain order, such as going to a restaurant. We all know, for example, that people enter a restaurant before they eat the food. An organizing script structure enables script items to be processed more efficiently, and therefore differently from non-script items. In addition it can be assumed that these processing differences have a differential effect on memory and are demonstrated by differential retrieval of typical and atypical script elements.

Differential retrieval of script and non-script elements is addressed by the script pointer plus tag

hypothesis of Graesser, Gordon, and Sawyer (1979). In this model, every event that is encoded, for which there exists a prior knowledge structure, results in a specific memory representation that consists of a pointer to the appropriate script and tags to items that are atypical; that is, unrelated or inconsistent with the script. Two predictions follow from this hypothesis. The first prediction is that memory discrimination will be better for atypical items than for typical script items. Discrimination entails correctly identifying that an item was present when it was and correctly rejecting an item that was not present. Thus, there should be a greater difference between correct identifications and correct rejections for atypical items than the difference between correct identifications and correct rejections for typical items. The second prediction is that there will be no memory discrimination for the occurrence/nonoccurrence of items that are very typical of the script. Because very typical items are part of the script they are not specifically stored. Thus, according to the model, it should be impossible to distinguish between typical items that are part of the script but not in a particular presentation and very typical items that were in the presentation.

Graesser, Gordon, & Sawyer (1979) supported both of the predictions from their theory. Participants first read a passage about a frequently enacted activity. Following a

one hour interval, recognition hits and false alarms were examined for items in the passage. A hit was the correct recognition of a sentence stated in the story. Similarly, a false alarm was the incorrect recognition of a sentence not stated in the story. False alarm rates were significantly higher for very typical items than for atypical items. Graesser et. al. observed that typicality had a far greater effect on false alarm rates than hit rates and therefore concluded that discriminative memory was better for atypical than for typical items. Their second prediction, that participants would be unable to distinguish between highly typical items that they had heard previously and new highly typical items, was also supported. They found that the mean hit rate for very typical items did not differ significantly from the mean false alarm rates for the same items. In addition, the prediction that discriminative accuracy is greater for atypical items than for typical items was further confirmed by Bower, Black, and Turner (1979).

Evidence supporting a differential decay rate for recall and recognition measures has clouded the issue. Smith and Graesser (1981) examined memory for scripted passages, measuring both recall and recognition over retention intervals ranging from 30 minutes to 3 weeks for atypical and typical items. In their analysis, they used a memory score that controlled for guessing in recognition and recall. Smith and Graesser wanted a measure that resembled

the d' measure from Signal Detection Theory in order to compare memory discrimination for recognition to memory discrimination for recall. The d' measure can be calculated for recognition but not for recall (Locksley, Stangor, Hepburn, Grosovsky, and Hochstrasser, 1984). The recognition memory score was obtained by subtracting the probability of a false alarm from the probability of a hit and dividing by one minus the probability of a false alarm. The memory score for recall is similarly calculated and is the probability of recall minus the probability of an intrusion divided by one minus the probability of an intrusion. From these scores, Smith and Graesser concluded that both recognition and recall scores decreased over the retention interval and that recognition was better for atypical than for typical items. In addition, they observed that recall memory was initially better for atypical items but that over time there was greater memory loss for atypical items, and that recall was better for typical than for atypical items after 3 weeks. They concluded that the script was utilized differently at recall from the way it was utilized at recognition.

These conclusions are based on the use of the memory score, used to "correct for guessing" as a measure of discrimination. By using this measure, Smith and Graesser wanted to examine memory performance, or memory strength and not the decision process and response criterion.

Unfortunately, there are difficulties with examining only memory discrimination. Lockhart and Murdock (1970) gave a strong warning concerning the interpretation of the d' discrimination measure. They described a situation where the d' measure was greater for memory of low frequency words than for memory of high frequency words. They suggested that this should not be interpreted as greater memory strength for the low frequency words, rather it should be interpreted as a greater difference between the strength of old and new low frequency words than the difference between the strength of old and new high frequency words. Note that all yes responses, both hits and false alarms would be expected to be higher for high frequency words than for low frequency words. In one sense, memory strength, in terms of the probability of output, is greater for the high frequency words although memory discrimination, the ability to identify whether a word has occurred in a particular setting, is greater for low frequency words.

It is not at all clear that the memory score used by Smith and Graesser (1981) to assess discrimination provides an adequate measure of discrimination. The Smith and Graesser memory discrimination score is calculated by subtracting the probability of false alarms from the probability of hits and dividing that number by the probability of correct rejections. Using this formula, the same memory discrimination score can be obtained for a hit

rate of 1 and a false alarm rate of .9, relflecting a high rate of yes responses and fairly low discrimination (which results in a memory score of 1) and for a hit rate of 1 and a false alarm rate of 0, reflecting perfect discrimination, (which also results in a memory score of 1). From this example, it is clear that the memory score is somewhat less than adequate for assessing discrimination. Another method of assessing discrimination may prove more useful.

In examining memory for low and high typical items in a script I want to be able to assess both people's discrimination memory and the probability of their retrieving items. For this purpose it may be instructive to examine hit and false alarm measures. Locksley, Stangor, Hepburn, Grosovsky, and Hochstrasser (1984) point out, however, that hits and false alarms are not independent, both are affected by shifts in the response criterion. This should be borne in mind when interpreting the measures and the relationship of hits and false alarms should be considered.

I graphed the hits and false alarms from Smith and Graesser 's (1981) data and a different picture emerged (see Figure 1). From this figure, six conclusions can be tentatively drawn. First, there are more hits (both recognition and recall) for typical items than for atypical items. Second, there are more intrusions/false alarms for typical than for atypical items. Third, there are more hits

over all for recognition than for recall. Fourth, false alarms/intrusions increase over time. Fifth, recall and recognition hits decrease over time for both typical and atypical items and sixth, atypical items decay at a faster rate than typical items.

Insert Figure 1 about here

Looked at in this fashion, the pattern seems to be the same for both recognition and recall. It can be concluded that similar script processes are operating on recall and recognition. In both forms of retrieval, while hits are slightly higher for typical than atypical items, false alarms and intrusions are considerably higher for typical items. Thus, discriminative memory is lower for typical items than for atypical items for both recognition and recall.

Smith and Graesser (1981) failed to report the sufficient statistics necessary in order to perform the analyses to justify these conclusions. In addition, they had only 6 participants per condition and furthermore, there were only four measures per participant for each of the typical and atypical items. As a result, it is possible that their results would not be replicable. Therefore the

first goal of this study was to attempt to replicate Smith and Graesser's findings using both more participants and more measures per participant.

Much of the research on scripts has been conducted using passages about mundane activities, such as going to a restaurant. An additional issue of interest in this study was the extension of script research to affect laden activities. Therefore the study was conducted using two passages, one about going to a restaurant and the other about a marital dispute.

The second goal of this study was to examine script content. Previous studies have relied mainly on only two methods for determining script content: participant free generations of passages and participant ratings of the typicality of items believed to be script-typical (see Smith and Graesser, 1981, for an example of this methodology). In order to confirm predictions about script content made from using data from these two methodologies it is necessary to use a third method, in this case examining intrusions. It is reasonable to assume that actions that are part of a participant's script but are not stated in the original passage will intrude into retrieval as recognition false alarms and recall intrusions. In order to examine the intrusions, three stories about the marital dispute were constructed with portions omitted that are predicted to be part of the script on the basis of free generations and

typicality studies. It was predicted that the omitted portion would be intruded into recognition as false alarms.

STUDY 1

Pretest

60 men and 60 women enrolled in introductory psychology classes at the University of Alberta participated as an option for partial fulfillment of course credit. Each student was given 1 of 12 booklets. The booklets contained a short selection from either a story about a restaurant or a story about a marital dispute. The stories were missing either their beginning, their middle, or their end. Students were asked to rate a series of short sentences relating to the selection of the story that they did not read. For example, a student asked to rate sentences relating to the middle of the marital dispute, such as, "They yell at each other", would read a short paragraph about the beginning of the dispute and a short paragraph about the resolution of the dispute. This procedure allowed for instantiation of the schema by providing the framework of the story without overly biasing students' ratings. Students were first asked to rate the sentences for consistency with the story theme. Following this, students were asked to rate the sentences for the likelihood of occurring in the appropriate story. The ratings of the pilot study for the sentences that I used in constructing the stimulus material for study 1 and study 2 are detailed in Appendix A.

Mandler (1984) suggested a distinction between the probability of an action and the relevance of that action to the theme. Consequently, in this thesis typicality was defined as both high likelihood ratings and high consistency ratings.

Rationale

It was predicted that the pattern of hits and false alarms for typical and atypical sentences would be the same for both recognition and recall. Specifically that pattern, based on the data from Smith and Graesser (1981) is illustrated in Figure 2. The mean scores fall into three clusters: cluster one, typical-sentence-hit-all delays and typical-sentence-false alarm-40 minute delay; cluster two, atypical hits all delays and typical sentence 2 day delay; cluster three, atypical false alarm all delays. The typical sentence false alarm 1 week delay mean is somewhat isolated. This pattern is represented by the contrast weights in Figure 2.

Insert Figure 2 about here

Method

Subjects and Design. Sixty men and 60 women enrolled in introductory psychology classes at the University of

Alberta participated as an option for partial fulfillment of course credit. A 2 (men, women) x 3 (40-minute delay, 2-day delay, 1-week delay) x 2 (recall, recognition) x 2 (stated, unstated) mixed factorial design was employed . The first four factors were between-subjects and the last factor was within-subjects. With the exception of sex, participants were randomly assigned to condition.

Materials. Dobbs and Rule (1986) identified the elements in the marital dispute script using free generations. They asked participants to list the actions in a husband and wife dispute. They used the Bower, Black, and Turner (1979) criterion of 25% of participants listing an element to determine the categories of elements in the marital dispute script. These categories were used in this thesis to construct two versions of a story about a marital dispute. Both versions contained one specific instance representing each of the general script categories: dissension, frustration, emotional state, emotional expression, verbal aggression, physical aggression, aggression interrupter, rational thought, apology, and outcome. The stories were systematically constructed using the consistency and likelihood ratings obtained from students in the separate pretest-study. High typical sentences had both high likelihood and high consistency ratings. Similarly, low typical sentences had both low

likelihood and low consistency ratings. A high typical and a low typical sentence were selected for each category. Version A of the story contained highly typical sentences from the dissension, emotional states, emotional expression, rational thought and aggression interrupter categories and low typical sentences from the frustrator, verbal aggression, physical aggression, apology, and outcome categories. Version B, on the other hand, contained highly typical sentences from the frustrator, verbal aggression, physical aggression, apology, and outcome categories and low typical sentences from the dissension, emotional states, emotional expression, rational thought and aggression interrupter categories.

The order of the sentences used in this thesis was determined using the orders obtained from a study by Dobbs and Rule (1986). Dobbs and Rule gave participants the elements of the marital dispute script in random orders and asked the participants to arrange the elements in the order in which they normally occurred. They found that the elements in the marital dispute script were consistently ordered. This ordering was used in constructing the marital dispute stories listed in Appendix B, each containing five typical elements and five atypical elements. In addition, two versions of a story about an evening at a restaurant were constructed using the typicality ratings from the separate pretest described previously. Each version

contained six typical elements and five atypical elements (see Appendix B for the stories). For both the restaurant and marital dispute stories the size of the element was arbitrarily determined.

Procedure. All students were given a cover story: they were told that they were participating in a study on computer assisted instruction of reading skills. They were instructed to listen to stories presented on headphones, and to answer questions about the stories. Further, they were told that a second session was required for the presentation of stories on a computer terminal. This cover story provided a rationale for a second appointment. Following this introduction students heard four stories and answered questions regarding the presentation, readability and ease of comprehension of each story (see Appendix C for the questionnaire). The first and last stories were buffer stories about a camping trip and buying a dress. The middle two stories were the stories of interest and were the marital dispute story and the restaurant story. The order of the four stories was kept constant so that each student first heard the camping trip story, followed by a story about a marital dispute, a story about a restaurant, and finally, the story about the blue dress. Students heard either Version A or Version B of the marital dispute and the restaurant story. At the end of the first session students

were reminded to return for the second scheduled session involving the computer terminal. Students left for either a 40-minute interval, a 2-day interval or a 1-week interval. In order to keep the sessions as similar as possible those students asked to return after 40 minutes were not asked to perform a specific intervening task. Instead they were free to do course work, or go for coffee as they chose.

When students returned for the second session, they were instead handed a booklet. The booklet consisted of either a recognition task for the restaurant story and a recall task for the husband and wife dispute story, or a recognition task for the husband and wife dispute story and a recall task for the restaurant story. For the recognition task students were given all the sentences in version A and in version B and were required to respond "In the story" or "not in the story" for each sentence. In the recall task students were given the title of the story and asked to write down everything that they could remember in as much detail as possible. Following this, students were fully and sensitively debriefed, as well as thanked for their help.

The recognition tests consisted of all of the sentences in both version A and version B of the relevant story. Hit rates and false alarm rates were calculated for each student from the recognition tests. A hit was a "yes" response to a sentence from the version heard by the student. A false alarm was a "yes response" to a sentence from the version

not heard by the student. Using the method outlined in Smith and Graesser (1981), recall hits and recall false alarms were calculated from the free recall protocols. Recall hits were items that were judged by two independent raters to be similar in content to one of the items in the version of the passage that the student heard. False alarms were items that were judged to be similar to one of the items in the alternative version of the script. In addition, recall items that were only slightly related to the sentence from either version of the story were coded as "related". Recall items that were completely unrelated to sentences in the story were coded as "intrusions". Agreement of the raters prior to discussion was 89.72%; following discussion agreement was 100%.

Results

Is the pattern of recall scores similar to the pattern of recognition scores? To answer this question, four planned comparisons using the contrast weights were performed on the recognition scores and on the recall scores for both the marital dispute story and the restaurant story. The initial predictions for the planned comparisons did not include any differences between men and women and so the data were collapsed over sex. The story factor was treated as a fixed effect (see Wike and Church, 1976, for a discussion of this

issue). Further, the traditional alpha level of .05 was used.

For the restaurant story the mean percentage of hits and false alarms with the contrast weights are presented in Table 1. Examining this, it appears that the hits and false alarms for both recall and recognition closely follow the predicted pattern the of the contrast weights. To test this the contrast weights were applied to the means of the restaurant/recall conditions, $F(1, 57) = 95.1$, $p=0.00$. The hits and false alarms for the restaurant/recall conditions are plotted in Figure 3. The same contrast weights were applied to the means from the restaurant/recognition conditions, $F(1, 57) = 445.7$, $p=0.00$. The hits and false alarms for the restaurant/recognition conditions are plotted in Figure 4. For the restaurant story the same pattern of hits and false alarms is seen for recall as for recognition.

Table 1

The percentage of hits and false alarms for the restaurant story with contrast weights

Condition	Weights	Retrieval: Recall	Retrieval: Recognition
30 Minutes			
Hits			
Typical	+18	27.5	70.83
Atypical	+18	39.0	69.0
False alarms			
Typical	+ 9	12.5	23.33
Atypical	-35	1.0	11.0
2 days			
Hits			
Typical	+18	18.34	65.0
Atypical	+ 9	16.0	59.0
False alarms			
Typical	+ 9	8.34	30.0
Atypical	-35	.01	15.0
1 week			
Hits			
Typical	+18	23.34	60.83
Atypical	- 3	11.0	45.0
False alarms			
Typical	+ 9	6.51	39.17
Atypical	-35	0.0	16.0

Insert Figure 3 about here

Insert Figure 4 about here

A similar finding occurred for the marital dispute story. Table 2 presents the mean percentage of hits and false alarms for the marital dispute story with the contrast weights. Again, it appears that the hits and false alarms for both recall and recognition closely follow the predicted pattern of the contrast weights. The planned contrasts were applied to means from the marital dispute/recall conditions, $F(1, 56) = 34.61$, $p=0.00$ and to the means from the marital dispute/recognition conditions, $F(1, 57) = 226.2$, $p=0.00$. The percentage of hits and false alarms for the marital dispute/recall and the marital dispute/recognition conditions are plotted in Figure 5 and 6. In summary, the pattern of hits and false alarms for typical and atypical sentences was the same for both recall and recognition.

Table 2

The percentage of hits and false alarms for the marital dispute story with contrast weights

		Retrieval	
Condition	Weights	Recall	Recognition
30 Minutes			
Hits			
Typical	+18	9.0	73.0
Atypical	+18	13.0	79.0
False alarms			
Typical	+ 9	2.0	36.0
Atypical	-35	0.0	11.0
Two days			
Hits			
Typical	+18	3.2	78.0
Atypical	+ 9	5.2	72.0
False alarms			
Typical	+ 9	2.2	51.0
Atypical	-35	1.0	21.0
One week			
Hits			
Typical	+18	5.0	71.0
Atypical	- 3	2.0	67.0
False alarms			
Typical	+ 9	2.0	44.0
Atypical	-35	0.0	28.0

Insert Figure 5 about here

Insert Figure 6 about here

To determine whether there were any effects for sex of the student, and to examine any absolute differences between recall and recognition for the restaurant and marital dispute data an analysis of variance, with 2(sex) x 2(mode of retrieval) x 2(typicality) x 2(stated, unstated) was performed. The first factor was between subjects, and the last three factors were assigned within subjects. This analysis revealed that there was no main effect for sex, $F(1, 117) = .35$, ns, $p=.557$. There was, however, a two-way interaction between sex and hits/false alarms, $F(1, 117) = 7.5$, $p=.007$ which was modified by a significant three way interaction among sex, mode of retrieval and hits/false alarms, $F(1, 117) = 7.36$, $p=.008$. Examining the mean hits and false alarms in Table 3 and the post-hoc Newman-Keuls revealed that women were more accurate than men on the recognition task but did not differ from men on the recall task. From this it appears that the women had greater discriminative accuracy for recognition than the men; their responses were characterized by a larger difference between hits and false alarms than the difference between hits and false alarms for the men.

Table 3

The percentage of hits and false alarms by men and women

	Men	Women
Hits		
recall	13.90b	16.0 b
recognition	63.81e	71.05f
False alarms		
recall	3.62a	4.0 a
recognition	30.48c	23.81d

Note. Means with different subscripts differ at .05 or less

Overall, there was a main effect for mode of retrieval, $F(1, 117) = 735.25$, $p=.001$. There were more yes responses for the recognition task than there were items recalled on the recall task. In addition there was a main effect for hits/false alarms, $F(1, 117) = 283.49$, $p=.001$. and for typicality, $F(1, 117) = 30.59$, $p=.001$. There were more yes responses to items that were in the story that students had heard and more yes responses for typical items. These conclusions are qualified by three two-way interactions and a three-way interaction. There was a significant interaction between mode of retrieval and hits/false alarms, $F(1, 117) = 143.34$, $p=.001$. It appears that discriminative

accuracy was greater for recognition than for recall; there was a larger difference between hits and false alarms for recognition than the difference between hits and false alarms for recall. Further, there was an interaction between the mode of retrieval and typicality, $F(1,117) = 17.05$, $p=.001$. It appears that there was a greater difference between "yes" responses to high typical items and low typical items for recognition than for recall. The two-way interaction between hits/false alarms and typicality, $F(1, 117) = 42.58$, $p=.001$, revealed greater discriminative accuracy for low typical than high typical sentences; there was a larger difference between hits and false alarms for low typical sentences than the difference between hits and false alarms for high typical sentences.

These two-way interactions were qualified by the three-way interaction between mode of retrieval, hits/false alarms, and typicality. Examining the three-way interaction gives a fuller picture of the relationship among these variables. The mean hits and false alarms for this three-way interaction are presented in Table 4. This interaction was significant, $F (1, 117) = 13.16$, $p=.001$. A Newman-keuls analysis revealed that there is somewhat greater discriminative accuracy for low typical items than for high typical items in recognition.

There is a similar pattern for recall; however the discriminative accuracy is somewhat lower than for

recognition. Overall, there are more "yes" responses for high typical items than for low typical items with the exception of recall hits, whereas the means do not differ significantly for low and high typical items. In summary, the three-way interaction indicated that discriminative accuracy was greater for low typical than for high typical elements for both recall and recognition although discriminative accuracy was lower overall for recall than for recognition.

Table 4

The percentage of hits and false alarms by retrieval

	Retrieval	
	Recall	Recognition
Hits		
Typical	14.48a	69.2b
Atypical	14.46a	65.22c
False Alarms		
Typical	31.7e	40.16d
Atypical	0.68f	16.4a

In order to make a further comparison to Smith and Graesser (1981), the data were collapsed over the two

stories and memory scores were calculated for recognition and recall using the formula (probability of hits - probability of false alarms)/(1 - probability of false alarms), discussed earlier. The means for the memory scores by retention interval and typicality can be seen in Table 5. The trends are remarkably similar to those obtained by Smith and Graesser. In both my study and that of Smith and Graesser recognition memory scores appear to be higher than recall memory scores. Further, recognition memory scores appear higher for atypical than for typical items. In the case of recall, the scores appear initially higher for atypical items. At delayed recall there appears to be a cross over and memory scores appear somewhat higher for typical items. Finally, memory scores appear to decrease over time.

Unlike Smith and Graesser who calculated the memory scores only for means I calculated scores for each participant and this permitted statistical analysis of the scores. An analysis of variance, with 3(delay interval) x 2(mode of retrieval) x 2(typicality) was performed. The first factor was between subjects, and the last two factors were assigned within subjects. This analysis revealed a main effect for retrieval, $F(1, 114) = 161.40$, $p=.001$; memory scores were higher for recognition than for recall. In addition there was a main effect for delay interval, $F(2, 114) = 6.77$, $p=.002$ and a main effect for typicality,

$F(1, 114) = 8.77$, $p=.004$. The last two main effects were qualified by a two-way interaction between interval and typicality, $F(2, 114) = 3.79$, $p=.026$. A Newman-Keuls analysis revealed that memory scores were higher only for atypical items at 30 minutes. There were no significant differences among the memory scores for the other conditions. Memory scores for atypical items decreased over time; there was no significant decrease for typical items. In summary, the pattern of memory scores appeared to be similar to the pattern of scores that Smith and Graesser based their conclusions upon although most of the differences were not statistically reliable.

Table 5

Memory scores for typical and atypical items in recall and recognition over time.

	Retrieval	
Condition	Recall	Recognition
30 Minutes		
Typical	0.102	0.509
Atypical	0.259	0.695
Two days		
Typical	0.066	0.487
Atypical	0.104	0.579
One week		
Typical	0.081	0.389
Atypical	0.065	0.398

Recall order. Most of the recalls followed the same order as the original story, usually with some omissions. An example of this type of recall is provided in Appendix D. A minority, however, did not follow this pattern. Instead students either provided a summary or listed items in almost

random order, seemingly as an afterthought. Students who included a summary provided an initial overview of the outcome of the event and proceeded to go into more detail with the ordered story. Other students wrote down a series of ordered items and followed this with a list of unordered items suggesting that the unordered items were an afterthought. Examples of both of these recalls are provided in Appendix E.

STUDY 2

In attempting to extend a concept to a new domain it is important to determine whether the concept is appropriate. In this particular case it is important to determine the psychological reality of the script concept for people's knowledge about aggression. The purpose of study 2 was to provide further verification of the content of the marital dispute script using a different paradigm. The two methods previously used to determine the content of the marital dispute were participant free generations (Dobbs & Rule, 1986) and typicality ratings, obtained in this thesis. A third method was necessary to test predictions of script content based on the earlier methodologies. It was predicted that students who heard a marital dispute story with emotional states, emotional expressions, or physical aggression omitted would give more false alarm responses on those items in a recognition test. Recall measures were not used in study 2. The results of study 1 indicated that recognition scores are sufficient when examining the pattern of effects of typicality. Specifically, it was predicted that the number of yes responses (both false alarms and hits) for these categories would be the same for those students who had heard these categories as the number of yes responses for those students who had not heard these categories.

Method

Subjects and design. Thirty men and 30 women from introductory psychology courses at the University of Alberta participated as an option for partial fulfillment of course credit. A 3 (emotional states omitted, emotional expressions omitted, physical aggression omitted) x 2 (men, women) x 2 (version A, version B) x 2 (typicality) mixed design was employed. The first three factors were between subjects, and the last factor was within subjects.

Materials. The marital dispute story from study 1 was modified and used in study 2. For the "emotional states omitted" condition, the emotional states were removed from version A and from version B. The rest of the story, however, remained the same as in study 1: both versions contained one instance of dissension, frustrator, emotional expression, verbal aggression, physical aggression, aggression interruptor, apology, and outcome. For the "emotional expression omitted" condition the emotional expressions were omitted from the story. There were 10 stories, one version of the camping trip story, two versions each of the marital dispute story with emotional states omitted, emotional expressions omitted, and with the physical aggression omitted, two versions of the restaurant story and one version of the story about buying a dress.

Procedure. Students were given the same cover story as in study 1. They then heard four stories; the camping trip, the marital dispute story with portions omitted, the restaurant story and the story about buying a dress. After each story, students answered questions regarding the presentation and ease of comprehension of the stories. As in study 1 students were asked to return after 40 minutes and were free to do course work or go for coffee. Students returned after a 40 minute interval and were given two recognition tests. The first test contained all the sentences from version A and B of the complete marital dispute story. The second test contained all the sentences from version A and B of the restaurant story. Subsequently, the students were fully debriefed.

Results

Did the omitted emotional states, emotional expressions and physical aggression intrude at retrieval? Three chi-square analyses on the frequency of the number of subjects having one or more "yes" responses were performed for each selected category of emotional states, emotional expressions, and physical aggression. Typically, researchers are trying to show that the null hypothesis is false and are thus most concerned with type I error. Type I error is the probability of rejecting the null hypothesis when in fact, it is true. Traditionally, the alpha level is

set at .05. In this study I am trying to show that the null hypothesis is true and I am most concerned with type II error. Type II error is the probability of accepting the null hypothesis when in fact, it is false. Thus, in order to reduce type II error the alpha level was set at .30. For emotional states the chi-square analysis indicated that students who had not heard the emotions responded 'yes' as frequently as students who had heard the emotions, $\underline{X}$ (2) =1.294 ns, $\underline{p}$>.30. For physical aggression the chi-square analysis indicated that students who had not heard the physical aggression responded 'yes' as frequently as students who had heard the physical aggression, $\underline{X}$ (2) =2.35 ns, $\underline{p}$>.30. For emotional expressions, however,the chi-square analysis indicated that students who had not heard the physical aggression did not respond 'yes' as frequently as students who had heard the emotional expressions, $\underline{X}$ (2) =5.49 $\underline{p}$ < .10. See Table 6 for the frequencies of students with yes responses to one or more items from each of the critical categories of emotional states, emotional expressions, and physical aggression. The critical omitted categories lie along the diagonal. The number of yes responses to emotional states do not differ for those students that heard emotional states and and those students that did not hear emotional states. Nor do the number of yes responses differ for physical aggression. They do, however, differ for emotional expression.

Table 6

Frequencies of students with "yes" responses to one or more emotional states, emotional expression, physical aggression

	Category omitted from the story		
Category of recognition items	Emotional states	Emotional expressions	Physical aggression
Emotional States	17	19	17
Emotional Expression	19	14	18
Physical Aggression	18	18	15

A further series of chi-square analyses was conducted for typical items and atypical items separately. For the typical emotional state items the chi-square analysis indicated that students who had not heard emotional states responded 'yes' to typical emotional states as frequently as students who had heard the emotional states, $\underline{X}$ (2) = 2.186, ns, $\underline{p}$>.30. This pattern was repeated for emotional

expressions and physical aggression, for emotional expressions, $\underline{X}$ (2) = .476, ns, $\underline{p}$ > .30, and for physical aggressions, $\underline{X}$ (2) = 1.6, ns, $\underline{p}$ > .30. There were, however, significant differences between the 'yes' responses for students who had heard the atypical items and students who had not heard the items from the critical category. The differences were significant for emotional states, $\underline{X}$ (2) = 2.679, $\underline{p}$ < .30, emotional expressions, $\underline{X}$ (2) = 4.23, $\underline{p}$ < .30, and for physical aggression, $\underline{X}$ (2) = 8.602, $\underline{p}$ < .30. See Table 7 for the frequencies. Again, the critical omitted categories lie along the diagonal. It is apparent that there were no differences in the responses for typical elements. There were, however, differences in the responses to atypical items.

Table 7
Frequencies of students with "yes" responses for typical and atypical emotional states, emotional expression, physical aggressions.

Category of recognition items	Category omitted from the story: Emotional states	Emotional expressions	Physical aggression
		Typical	
Emotional States	14 -	17 (9)	11 (8)
Emotional Expression	14 (9)	13 -	15 (9)
Physical Aggression	10 (7)	8 (7)	12 -
		Atypical	
Emotional States	5 -	8 (6)	10 (8)
Emotional Expression	6 (5)	1 -	4 (3)
Physical Aggression	8 (5)	12 (9)	3 -

Note. Students who did not hear the critical category have no hits for that category. This is indicated by a dashed line. For the students that did hear the critical category hits are indicated in brackets.

A further series of chi-square analyses was conducted to compare the false alarm responses to typical items of students who heard the story with the critical category omitted with the responses of students who had heard only an atypical item in the critical category. The frequency of false alarms is presented in Table 8. No predictions were made for this analysis so the alpha level was set at the

usual .05. For emotional states and emotional expressions there were no significant differences, $\underline{X}$ (1) = .96, $\underline{p} > .05$, and $\underline{X}$ (1) = .417 $\underline{p} > .05$, respectively. There was, however, a significant difference for physical aggression, $\underline{X}$ (1) = 7.08, $\underline{p} < .05$. In Table 8 it can be seen that false alarms to typical emotional states do not differ between those students who did not hear any emotional states and those students who heard only atypical emotional states. This was the same for emotional expressions. For physical aggression there are far fewer false alarms to typical items for those students who heard atypical items only.

Table 8

Frequencies of false alarm responses for typical emotional states, emotional expression, physical aggressions.

Category of typical items	Elements heard by students from the critical category: No elements; Critical category omitted	Atypical elements only
Emotional States	14	11
Emotional Expression	13	11
Physical Aggression	12	4

There were no differences in hit rates for typical items among the three conditions, $\underline{X}$ (2) = 2.89, $\underline{p} > .05$. Further, there were no significant differences in the hit rates for atypical items, $\underline{X}$ (2) = 5.0, $\underline{p} > .05$.

GENERAL DISCUSSION

The pattern of effects for typicality ratings was the same for recall as was the pattern for recognition, even though the absolute levels of hits and false alarms were different. Recognition scores were higher than recall scores. The pattern of effects of typicality was predicted from the examination of hits and false alarms for typical and atypical items in Smith and Graesser (1981). This pattern of results was predicted as the outcome of script operation. Thus it appears that the schema operates at retrieval in the same way for recall as it operates for recognition. This is not to say that recall and recognition involve identical processes. Indeed, in this study there were higher recognition scores than recall scores, supporting a difference between recall and recognition. Mandler (1980) has proposed that a different type of information is retrieved in recognition than in recall. Recognition is said to involve both a familiarity judgement and a retrieval process based on an organizational search. Recall is said to involve only the retrieval process (see Mandler, 1980, for a further discussion). Presumably, it is in this retrieval process that the script is utilized, thus, both recognition and recall may utilize the script in a similar fashion.

This is a different conclusion from that made by Smith and Graesser (1981). Smith and Graesser suggested that the

script is utilized differently in recall from recognition. Their interpretations were based upon their formula for memory scores. However, I hesitate to interpret the results using memory scores in terms of discrimination because of the inadequacy of the memory scores as discussed previously. The data from Smith and Graesser were replicated in this thesis, however , their conclusions, based on the their manipulation of their data, were not supported. Examining hits and false alarms as was done in this thesis allows for interpretations that consider the pattern of results: both the absolute level of scores and discrimination among scores.

On a more pragmatic note, the similar utilization of the script at recall and recognition has implications for researchers interested in script content. The collection of both recognition measures and the more time consuming recall measures will be no longer necessary. Examining the pattern of effects for typicality for recognition will provide the same information as coding recall hits and intrusions.

As in other studies (Graesser, Woll, Kowalski & Smith, 1980, Nakamura, Graesser, Zimmerman, & Riha, 1985, Schmidt & Sherman, 1984), memory discrimination is greater for atypical items than for typical items. Further, hit rates were higher for typical items than for atypical items. This finding is consistent with the script pointer plus tag hypothesis that the memory representation consists of

atypical items that are tagged specifically and of a pointer to the script that contains typical items. Some of the typical items will be the same as those stated in the text, whereas the others will be inferred.

In addition, the pattern of effects of typicality were exhibited for the retrieval of the restaurant story and for the retrieval of the marital dispute story. Thus, it appears that the script concept can be extended to emotionally-laden events.

An unexpected finding was that women exhibited better discriminative memory at recognition than men. The women did not differ from the men at recall. Of note here, there were no sex differences for the typicality ratings from the pilot work.

The second focus of this study was to examine script content. It was determined that when emotional states, emotional expressions, and physical aggression were omitted from the story they later intruded at retrieval in the form of recognition false alarms for typical items. For emotional states and emotional expressions, false alarms to typical items were as frequent for omitted categories as for categories with only an atypical item. For physical aggression, however, there were more false alarms to typical items when physical aggression was omitted from the story than when an atypical physical aggression was present in the story. False alarms were not made to typical physical

aggression when the atypical physical aggression was heard. Students who heard the atypical physical aggression seemed to know that the typical physical aggression had not occurred. Thus they did not false alarm to the typical physical aggression; they either correctly recognized the atypical physical aggression or they rejected all physical aggression. Of note here, in the pilot study only 2 out of 44 physical aggression items were rated as typical for a marital dispute. In addition, both items ("Mark grabs Jean", "Mark taps Jean on the shoulder") were low intensity physical aggression. They were not particularly violent. The atypical item, ("Mark yanks Jean") was selected to match with the typical item in terms of specificity and intensity. This was accomplished by consultation with colleagues and not by the use of ratings. It is possible that "yanks" is more violent than "grabs". Further, students who heard the atypical "yanks" may have thought that a more violent episode had occurred. Indeed, a script for a violent marital dispute may have been instantiated. Thus, students either would correctly recognize the atypical physical aggression or would look for a violent physical aggression typical for the violent marital dispute. This item would not be present in the recognition test and students would simply not identify any of the physical aggression items.

It would be interesting to examine responses to a recognition test with a wider range of physical aggression

items. False alarms to more violent physical aggression by those students who had heard the atypical "yanks" would be consistent with the hypothesis of activation of a more violent script. Further it would be interesting to compare typicality ratings for a variety of marital disputes of different intensity. I would expect that the typicality of physical aggression elements would change with the intensity of the dispute. For example, "beating" would probably be typical of a violent dispute but would probably not be typical of a less intense dispute.

Bower, Black and Turner, (1979) proposed that "the thoughts and feelings of the character ... have no essential place in the causal flow of events" and that these feelings are "irrelevant statements". This study demonstrated, however, that these categories are an integral part of people's knowledge about marital disputes. Indeed, emotions may play an important role in the causal chain of events. For example, characters may be aggressive because of their emotions. Further studies are necessary to investigate the role of emotions in the causal chain of events. At present, it appears that the script concept can be extended to comprehension of aggressive incidents. Caution should be taken in extending the script concept to all emotionally-laden events without further research because there is some evidence that positive emotions are processed

differently form negative emotions (Fiske & Taylor, 1984; Isen, 1985).

Script theories are a subset of schema theories. Schema theories in general have been criticized for being imprecise and for producing contradictory predictions (Fiske & Linville, 1980). The results across a number of studies, however, have been consistent (Graesser, Woll, Kowalski & Smith, 1980, Light & Anderson, 1983, Nakumura, Graesser, Zimmerman, & Riha, 1985, Schmidt & Sherman, 1984, Smith & Graesser, 1981). These studies include different retention intervals, different subject matter, and different ages of subjects. These studies show that there is greater discriminative accuracy for atypical script material than for typical script material. In this thesis this conclusion was extended to include explicitly-emotional material.

Schema theorists have demonstrated consistent results thus overcoming the arguments of their critics. At present, however, a behavioural component to schema theories is lacking. Some future research avenues in this area could include examining subject responses to typical and atypical actions, exploring the development of social schemata and consequent changes in behaviour. For schema theories to be fruitful in advancing knowledge in social cognition the implications of schema theories for behaviour will have to be explored.

References

Bartlett, F. (1932). Remembering: A study in Experimental and social psychology. London: Cambridge University Press.

Bower, G., Black, J., & Turner, T. (1979). Scripts in memory for text. Cognitive Psychology, 11, 177-220.

Dobbs, A., & Rule, B. (1986). Scripts for aggression. Paper presented at the meetings of the Canadian Psychological Association, Toronto, June, 1986.

Fiske, S., & Linville, P. (1980). What does the schema concept buy us? Personality and social Psychology Bulletin, 6, 543-557.

Fiske, S., & Taylor, S. (1984). Social Cognition. Reading, Mass.: Addison-Wesley p326-327

Graesser, A., Woll, S., Kowalski, D., & Smith, D. (1980). Memory for typical and atypical actions in scripted activities. Human Learning and Memory, 6, 503-515.

Graesser, A., Gordon, S., & Sawyer, J. (1979). Recognition memory for typical and atypical actions in scripted activities: Tests of a script pointer + tag hypothesis. Journal of verbal learning and verbal behavior, 18, 319-332.

Isen, A. (1985). Asymmetry of happiness and sadness in effects on memory in normal college students: Comments on Hasher, Rose, Zacks, Sanft, and Doren. Journal of Experimental Psychology: General, 114, 388-391.

Knoke, D., & Burke, P. (1980). Log-linear models (Sage University Paper series on Quantitative Applications in the Social Sciences, 07-020). Beverly Hills and London: Sage Publications.

Light, L., & Anderson, P. (1983). Memory for scripts in young and older adults. Memory and Cognition, 1 435-444.

Lockhart, R., & Murdock, B. (1970). Memory and the theory of signal detection. Psychological Bulletin, 74, 100-109.

Locksley, A., Stangor, C., Hepburn, C., Grosovsky, E., & Hochstrasser, M. (1984). The ambiguity of recognition memory tests of schema theories. Cognitive Psychology, 16, 421-448.

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. Psychological Review, 87, 252-271.

Nakamura, G., Graesser, A., Zimmerman, J., & Riha,J. (1985). Script processing in a natural situation. Memory and Cognition, 13, 140-144.

Scmidt, D., Sherman, R. (1984). Memory for persuasive messages: A test of a schema-copy-plus-tag model. Journal of Personality and Social Psychology,

Smith, D., & Graesser, A. (1981). Memory for actions in scripted activities as a function of typicality, retention interval, and retrieval task. Memory and Cognition, 9, 550-559

Wike, E., & Church, J. (1976). Comments on Clarks "The language-as fixed-effect-fallacy". Journal of Verbal Learning and Verbal Behavior, 15, 249-257.

Appendix A

Pilot study ratings

The restaurant

	Consistency		Likelihood	
	Mean	Std. dev.	Mean	Std. dev.
Jack put away his tennis racket	4.2	0.919	4.2	0.789
Jack showered and got dressed	3.8	1.751	3.4	1.578
He put on a jacket	4.0	1.323	3.7	1.494
Jack and Chris let the valet park the car	3.5	1.716	3.4	1.897
Jack and Chris confirmed their reservations	1.6	0.843	1.6	0.966
The hostess escorted them to the table	1.0	0	1.5	1.269
They sat down and the hostess placed their menus on the table	1.2	0.422	1.4	0.966
They ordered drinks	1.7	1.16	1.5	0.707
Jack gave Chris a book he had borrowed	4.3	0.949	4.1	0.994
Jack straightened his collar	3.4	1.265	4.1	0.994
The waiter told them about the specials	1.2	0.422	1.5	0.707
They looked at their menus	1.2	0.422	1.4	0.966
They ordered dinner	1.2	0.632	1.4	0.699
They talked for a while	1.6	0.966	1.7	0.823

Jack put a pen back in his pocket	4.0	1.054	3.7	1.059
Jack picked a napkin off the floor	2.9	1.792	3.2	1.751
They ate their meal	1.5	0.972	1.5	0.527
They had dessert	1.1	0.316	1.6	1.265
Jack bought some mints	3.5	1.354	3.5	1.179
Jack cleaned his glasses	3.8	1.033	3.91	1.197
They walked out of the restaurant	1.1	0.316	1.5	1.269
Jack paid the bill	1.4	0.966	1.4	0.966

<u>The marital dispute</u>

	Consistency		Likelihood	
	Mean	Std. dev.	Mean	Std. dev.
Mark and Jean disagree about money matters	1.4	0.966	1.1	0.316
Mark and Jean disagree about vacation plans	3.4	1.174	3.2	1.476
They nag each other	2.4	1.174	2.4	1.265
They tease each other	3.5	1.354	3.3	1.252
They are annoyed	1.2	0.422	1.2	0.422
They are tired	3.8	1.549	3.7	1.567
They raise their voices	1.6	0.699	1.3	0.483
They whisper	4.9	0.316	4.6	0.516
They accuse each other	1.8	0.919	1.4	0.516
They lie to each other	4.0	0.943	3.6	0.843
Mark grabs Jean	2.2	1.317	1.9	0.994
Mark yanks Jean	3.6	1.075	3.8	1.398

They rationalize	2.6	1.35	2.4	1.35
They scheme	3.4	1.174	3.4	0.843
Jean avoids Mark	1.5	0.527	1.8	0.919
Jean makes coffee	4.2	0.919	4.1	1.287
They say they are sorry	2.3	1.252	1.5	0.527
They beg forgiveness	3.6	1.506	3.4	0.966
They find a resolution to their problem	2.7	0.949	1.8	0.919
They agree to disagree	3.8	1.033	3.6	1.075

Appendix B

Stimulus materials

The Camping Trip

Four friends, Jean, Karen, Debbie and Sue set off camping for three days, sharing the cost of gas. They had been to the campsite before and knew that it was a good place to go. When they arrived, they quickly set up camp. From their campsite they could hear the lake lapping against the shore. It was good for swimming because the water was waist-deep but it made their skin turn blue. There were many pebbles on the beach and the sand was coarse. During the day they set out for a hike. It was very hot and poor Sue came back with sore, red skin. Later in the evening they made a fire and roasted marshmallows. Everyone enjoyed the trip.

The Blue Dress

The dress in the window was a deep blue and gathered at the waist. Mrs. Kretch frequently walked past the shop when she went to get her groceries. The dresses on sale were much cheaper but none was as nice. One day Mrs. Kretch brought along her friend Emily in order to ask for her opinion. Emily agreed that it would be an extravagance but thought that the quality of the material was high. Furthermore, Emily thought that it was unlikely to go out of

style. Mrs. Kretch finished her regular grocery shopping. Then she decided to take the chance. She was very pleased with her decision. It really brought out the violet of her eyes.

A Marital dispute (Version A)

Mark and Jean disagree about money matters. First they tease each other and soon, both are annoyed. Mark and Jean raise their voices and lie to each other. Finally Mark yanks Jean. Following this, they rationalize and Jean avoids her husband. Later they beg forgiveness and agree to disagree.

A Marital dispute (Version A - no physical aggression)

Mark and Jean disagree about money matters. First they tease each other and soon, both are annoyed. Mark and Jean raise their voices and lie to each other. Following this, they rationalize and Jean avoids her husband. Later they beg forgiveness and agree to disagree.

A Marital dispute (Version A - no emotional states)

Mark and Jean disagree about money matters. First they tease each other. Mark and Jean raise their voices and lie

to each other. Finally Mark yanks Jean. Following this, they rationalize and Jean avoids her husband. Later they beg forgiveness and agree to disagree.

A Marital dispute (Version A - no emotional expression)

Mark and Jean disagree about money matters. First they tease each other and soon, both are annoyed. Mark and Jean lie to each other. Finally Mark yanks Jean. Following this, they rationalize and Jean avoids her husband. Later they beg forgiveness and agree to disagree.

A Marital dispute (Version B)

Mark and Jean disagree about vacation plans. First they nag each other and soon both are tired. Mark and Jean whisper and accuse each other. Finally Mark grabs Jean. Following this, they scheme and Jean makes coffee. Later they say they are sorry and find a resolution to their problem.

A Marital dispute (Version B - no physical aggression)

Mark and Jean disagree about vacation plans. First they nag each other and soon both are tired. Mark and Jean whisper and accuse each other. Following this, they scheme and Jean makes coffee. Later they say they are sorry and find a resolution to their problem.

A Marital dispute (Version B - no emotional states)

Mark and Jean disagree about vacation plans. First they nag each other. Mark and Jean whisper and accuse each other. Finally Mark grabs Jean. Following this, they scheme and Jean makes coffee. Later they say they are sorry and find a resolution to their problem.

A Marital dispute (Version B - no emotional expression)

Mark and Jean disagree about vacation plans. First they nag each other and soon both are tired. Mark and Jean accuse each other. Finally Mark grabs Jean. Following this, they scheme and Jean makes coffee. Later they say they are sorry and find a resolution to their problem.

The restaurant (Version A)

Jack put away his tennis racket. Next, he put on a jacket. Later, Jack and Chris confirmed their reservations with the hostess.

They sat down and the hostess placed their menus on the table. Jack gave Chris a book he had borrowed. After a while, the waiter told them about the specials and they ordered dinner. Jack put a pen back in his pocket before they ate their meal. At the end of the evening Jack bought some mints and walked out of the restaurant.

The restaurant (Version B)

Jack showered and got dressed. Next, Jack and Chris let the valet park the car. Later, the hostess escorted them to their table. They ordered drinks. Jack straightened his collar. After a while, they looked at their menus and they talked for a while. Jack picked a napkin off the floor before they had dessert. At the end of the evening, Jack cleaned his glasses and paid the bill.

Appendix C

Phony questionnaire

The story made a lot of sense to me

1	2	3	4	5
I completely agree	I agree somewhat		I disagree somewhat	I completely disagree

I found the passage easy to understand

1	2	3	4	5
I completely agree	I agree somewhat		I disagree somewhat	I completely disagree

I liked the method of presentation

1	2	3	4	5
I completely agree	I agree somewhat		I disagree somewhat	I completely disagree

Appendix D

A sample recall

Subject: 19. Male, 40-minute delay, heard version A

husband - Mark was his name (or maybe it was Jack)
wife - I can't remember hers.
they disagree about money
First they joke about it and eventually get on each others nerves.
then the husband gets angry with his wife. She is quiet or doesn't say much more.
Then they make up and agree to disagree.

Appendix E

Sample recalls

Subject: 42. Male, 1-week delay, heard version A

Two people enter a restaurant to buy something to eat. The two converse while being served their food. The two do not even talk about the food, they talk about other things.

They leave the restaurant at the end of the story when their conversation ends.

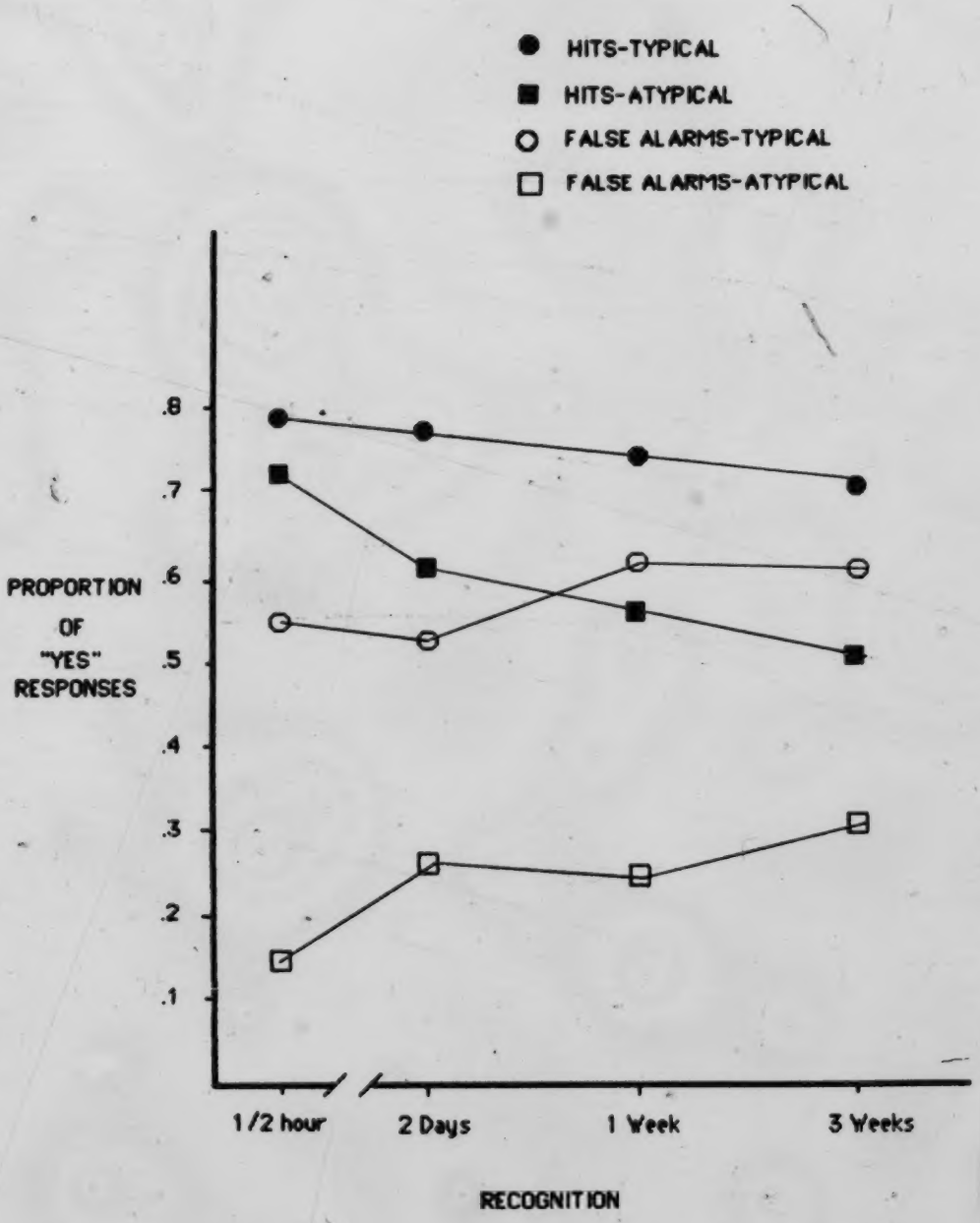
The two people ordered coffee first, then they ordered their meal.

Subject: 144. Female, 40-minute delay, heard version A

Mark and Jean are fighting over management of money matters. They, in the end, agree to disagree. Jean yells at Mark, and Mark yells back. But since they realize their problem, they are able to resolve it somewhat.

Figure Caption

Figure 1a. Recognition hits and false alarms data from Smith and Graesser (1981).


HITS-TYPICAL
HITS-ATYPICAL
FALSE ALARMS-TYPICAL
FALSE ALARMS-ATYPICAL
PROPORTION OF "YES" RESPONSES
.8
.7
.6
.5
.4
.3
.2
.1
1/2 hour
2 Days
1 Week
3 Weeks
RECOGNITION

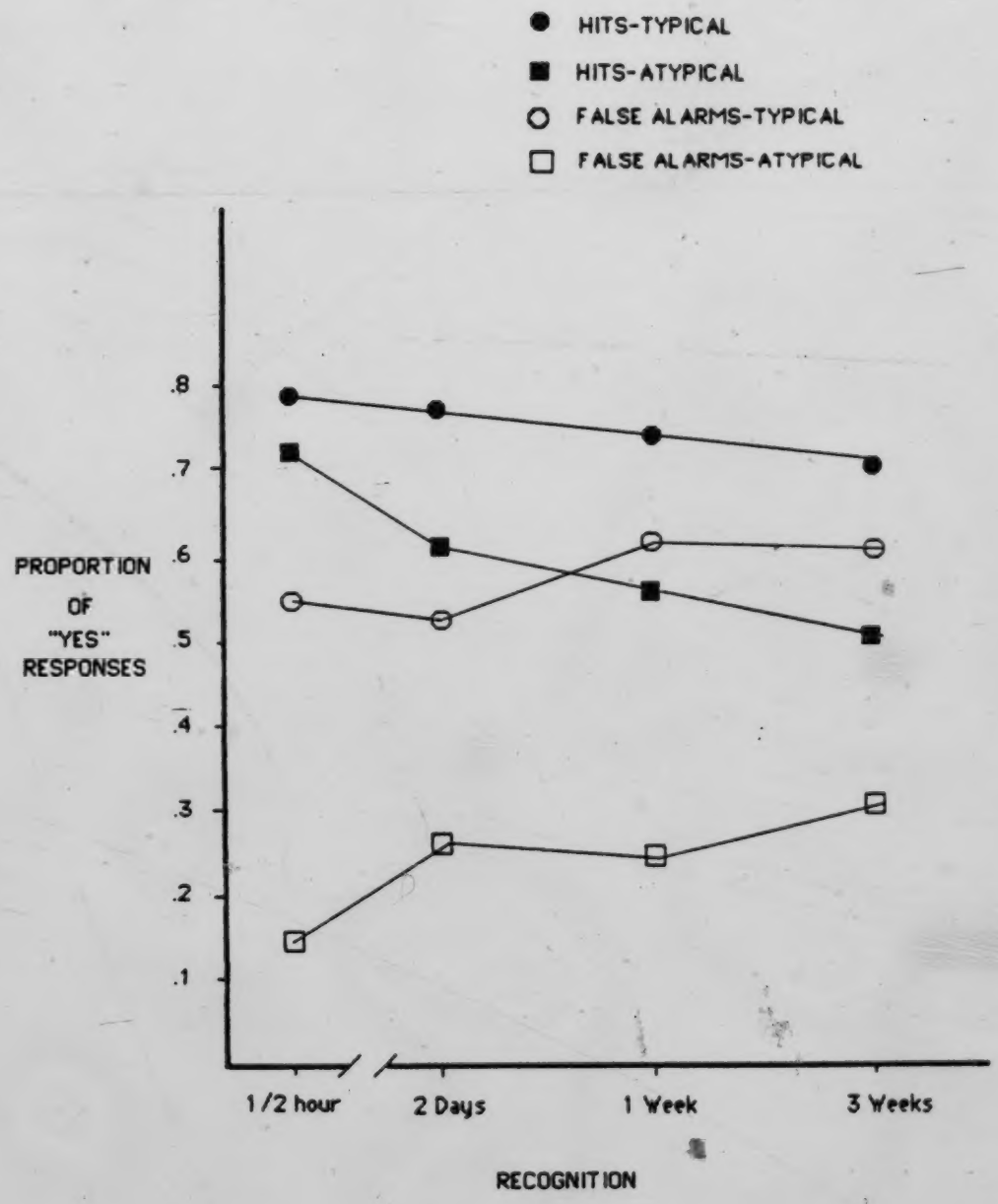

HITS-TYPICAL
HITS-ATYPICAL
FALSE ALARMS-TYPICAL
FALSE ALARMS-ATYPICAL
PROPORTION OF "YES" RESPONSES
.8
.7
.6
.5
.4
.3
.2
.1
1/2 hour
2 Days
1 Week
3 Weeks
RECOGNITION

Figure Caption

Figure 1b. Recall hits and false alarms data from Smith and Graesser (1981).

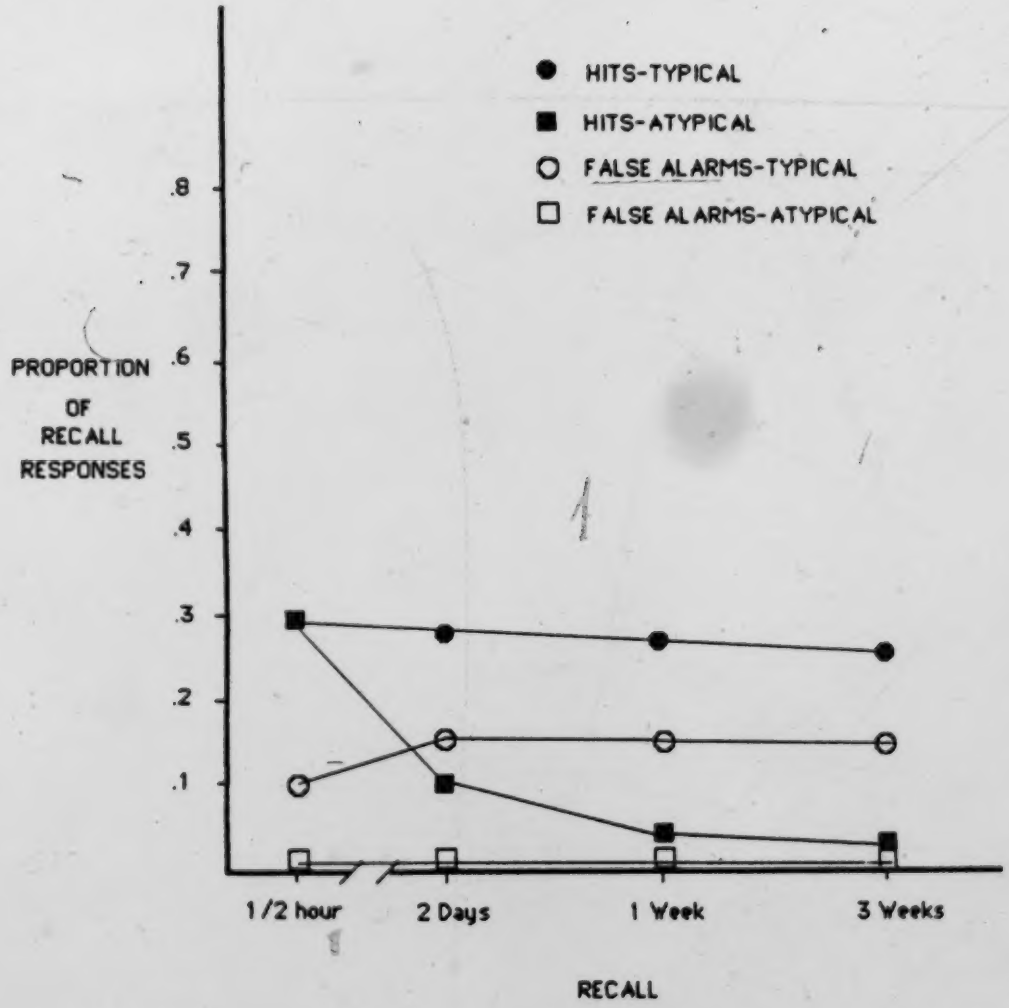

HITS-TYPICAL
HITS-ATYPICAL
FALSE ALARMS-TYPICAL
FALSE ALARMS-ATYPICAL
PROPORTION OF RECALL RESPONSES
.8
.7
.6
.5
.4
.3
.2
.1
1/2 hour
2 Days
1 Week
3 Weeks
RECALL

Figure Caption

Figure 2. Predicted pattern of hits and false alarms for recall and recognition.

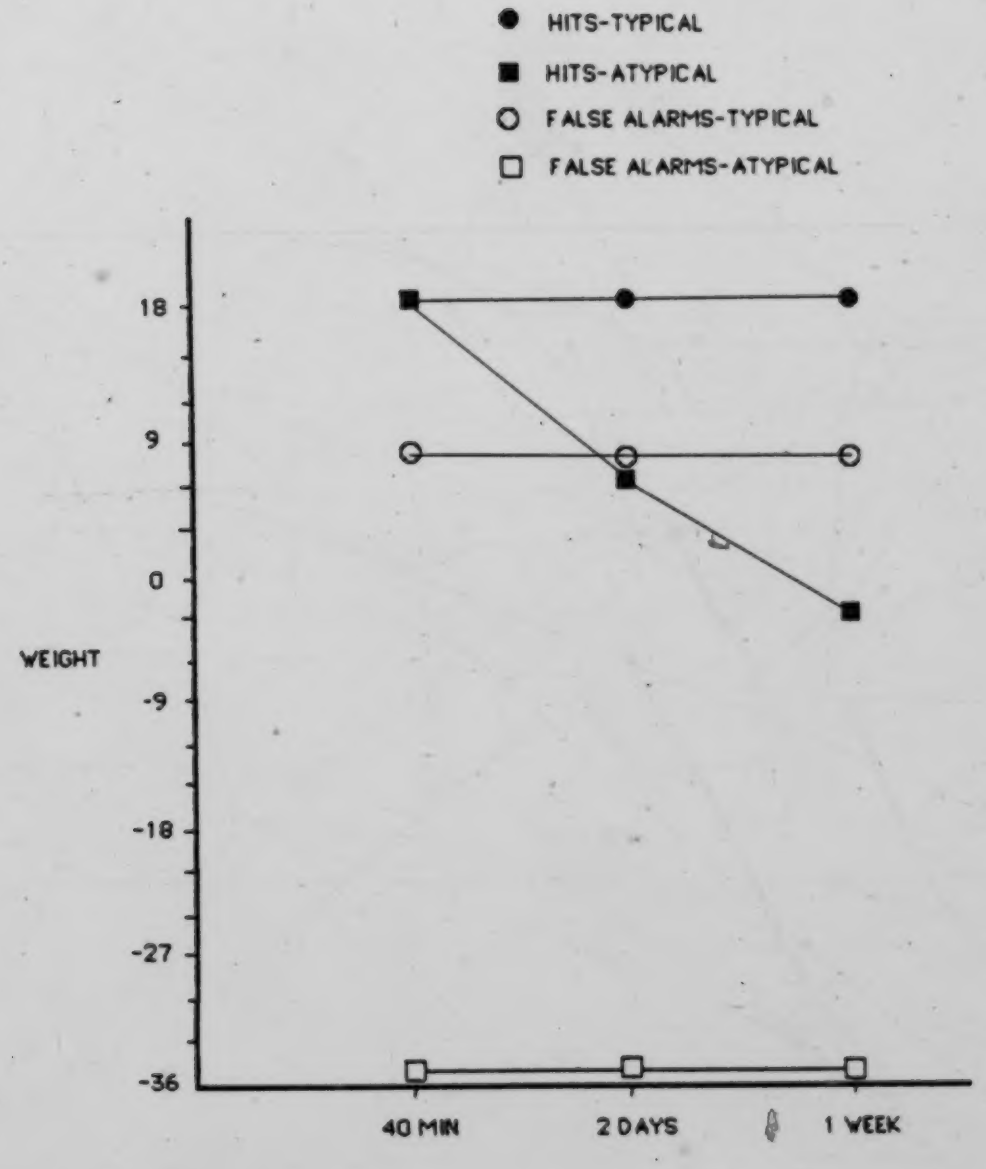

HITS-TYPICAL
HITS-ATYPICAL
FALSE ALARMS-TYPICAL
FALSE ALARMS-ATYPICAL
18
9
0
-9
-18
-27
-36
WEIGHT
40 MIN
2 DAYS
1 WEEK

Figure Caption

Figure 3. The percentage of hits and false alarms for the restaurant/recall conditions.

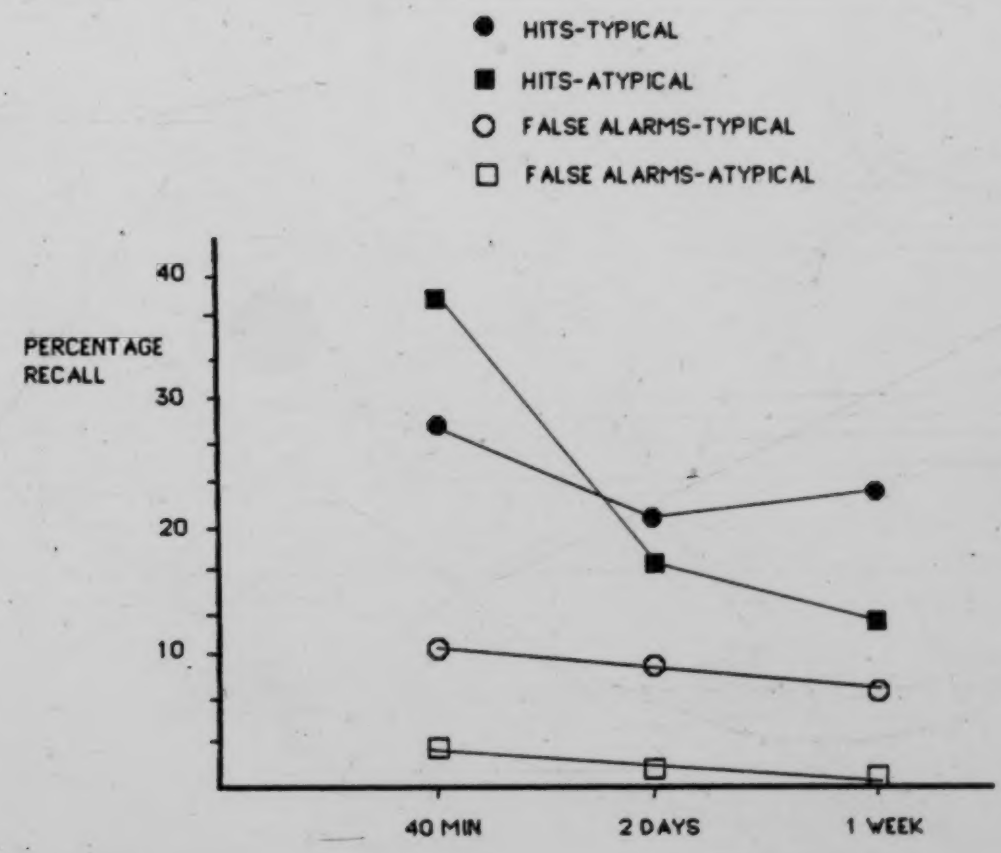

HITS-TYPICAL
HITS-ATYPICAL
FALSE ALARMS-TYPICAL
FALSE ALARMS-ATYPICAL
PERCENTAGE RECALL
40
30
20
10
40 MIN
2 DAYS
1 WEEK

Figure Caption

Figure 4. The percentage of hits and false alarms for the restaurant/recognition conditions.

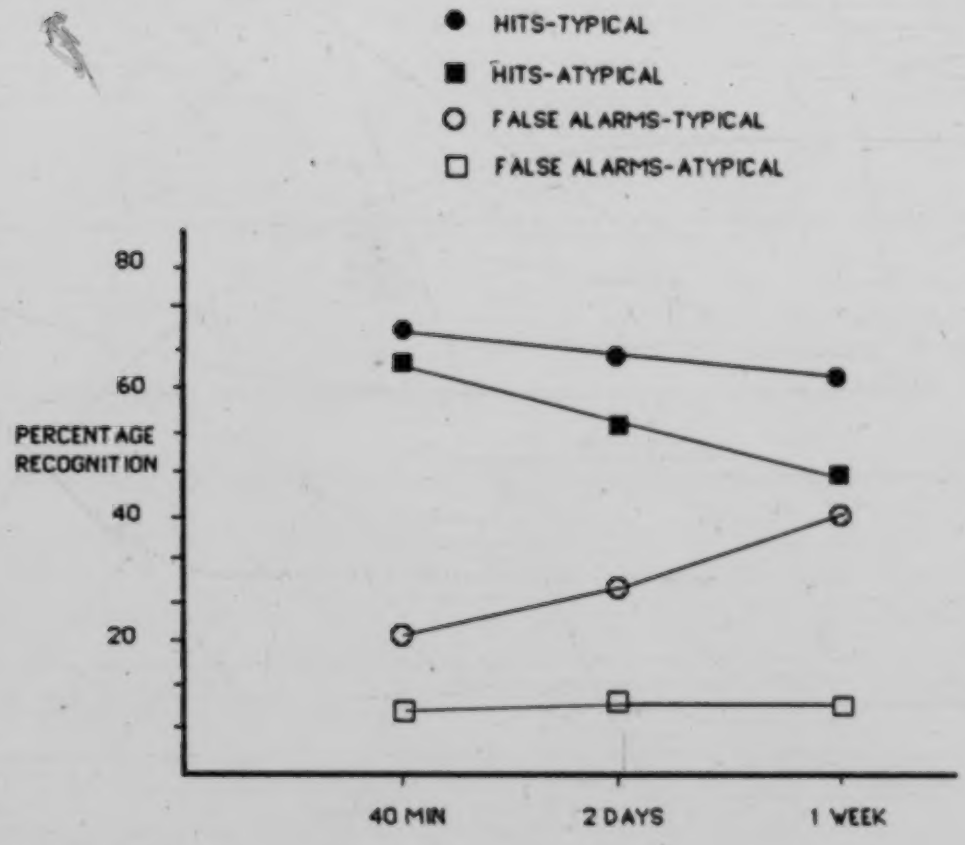

HITS-TYPICAL
HITS-ATYPICAL
FALSE ALARMS-TYPICAL
FALSE ALARMS-ATYPICAL
80
60
40
20
PERCENTAGE RECOGNITION
40 MIN
2 DAYS
1 WEEK

Figure Caption

Figure 5. The percentage of hits and false alarms for the marital dispute/recall conditions.

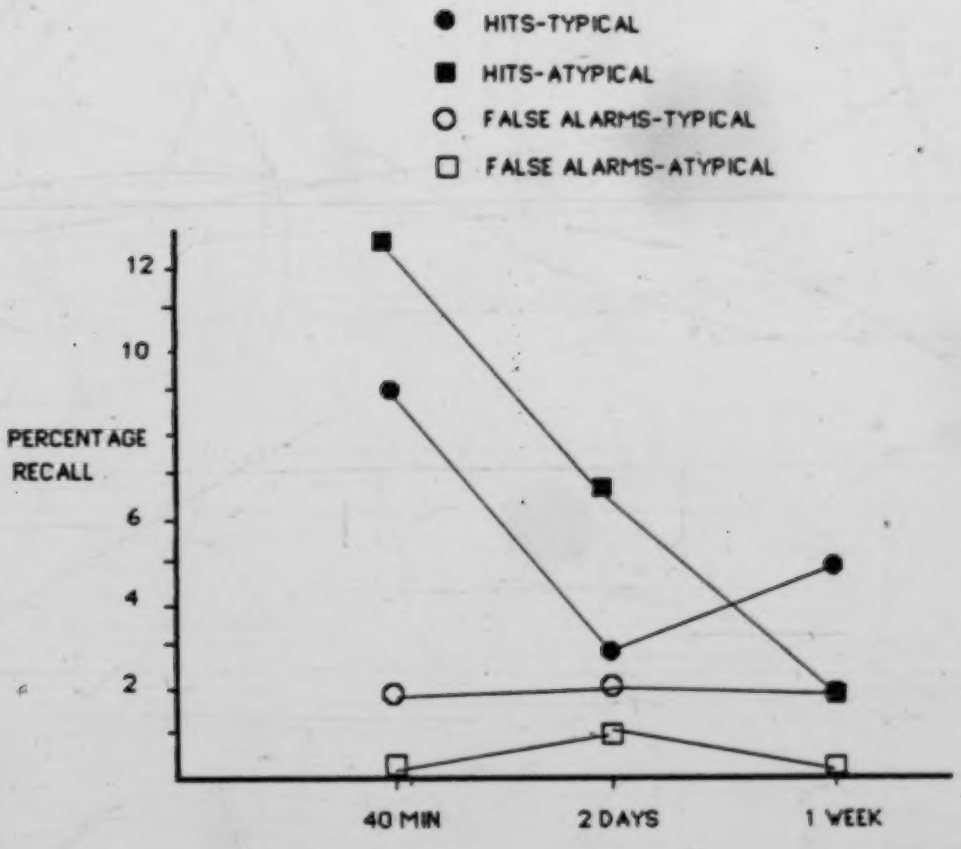
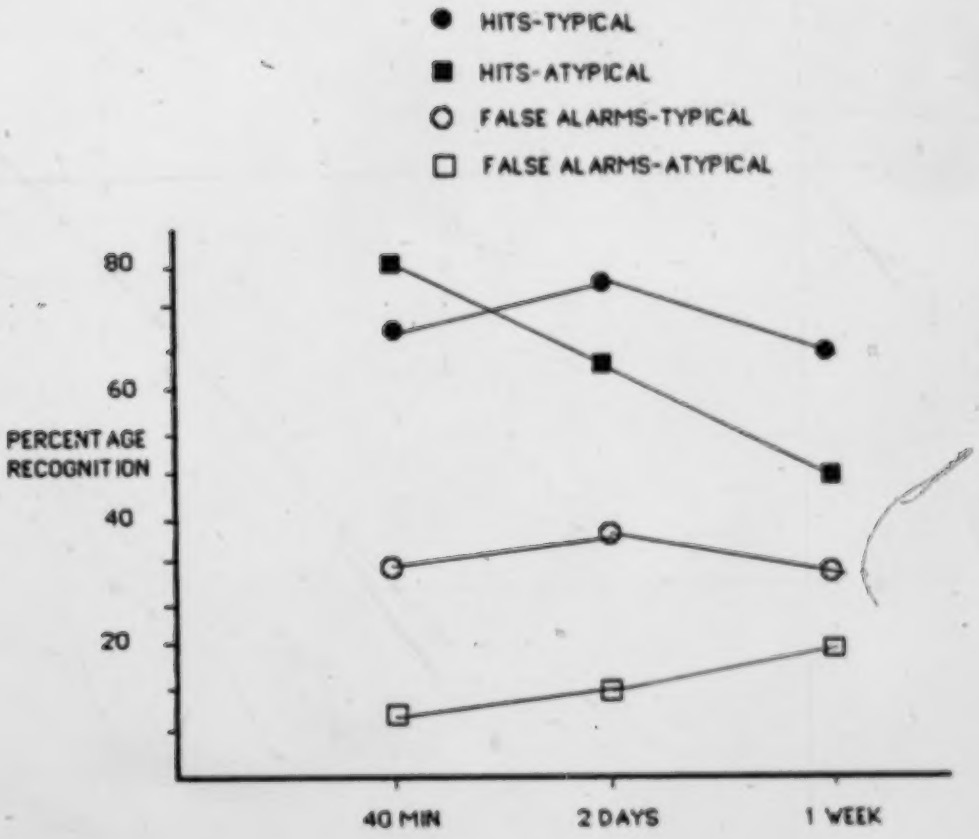

HITS-TYPICAL
HITS-ATYPICAL
FALSE ALARMS-TYPICAL
FALSE ALARMS-ATYPICAL
PERCENTAGE RECALL
12
10
6
4
2
40 MIN
2 DAYS
1 WEEK

Figure Caption

Figure 6. The percentage of hits and false alarms for the marital dispute/recognition conditions.


HITS-TYPICAL
HITS-ATYPICAL
FALSE ALARMS-TYPICAL
FALSE ALARMS-ATYPICAL
80
60
40
20
PERCENTAGE RECOGNITION
40 MIN
2 DAYS
1 WEEK